Exercício 5

P81

Pesos pesos $w \in \mathbb{R}$ ~~d~~

$$J(w) = \ell(w) + \lambda\left[\frac{1-\alpha}{2}\|w\|^2 + \alpha\|w\|_1\right],$$

$\lambda \geq 0,$

$\alpha \in [0,1]$

$\ell(w)$ é a perda, suave

a Regularização combina L2 e L1

Derivadas de cada termo

Como ℓ é suave

$$d\ell(w) = \left(\nabla \ell(w)\right)^T dw$$

$\boxed{\text{L2 (Ridge)}}$

$$RL_2(w) = \frac{\lambda(1-\alpha)}{2} w^T w \Rightarrow$$

$$dRL_2 = \lambda(1-\alpha) w^T dw \Rightarrow$$

$$\nabla RL_2(w) = \lambda(1-\alpha) w$$

exercicio 5 (182)

L1 (Lasso) $R_{L1}(w) = \lambda\alpha \sum_{j=1}^{d} |w_j|$

$|x|$ não é diferenciável em $x=0$

$$\partial|x| = \begin{cases} \{+1\}, & x>0 \\ [-1,1], & x=0 \\ \{-1\}, & x<0 \end{cases} \Rightarrow \partial R_{L1}(w) = \lambda\alpha \,\mathrm{sign}(w)$$

Subgradiente Total da Elastic Net

$$0 \in \nabla \ell(w) + \lambda(1-\alpha)w + \lambda\alpha \,\mathrm{sign}(w)$$

subgradiente de J é

$$\partial J(w) = \nabla \ell(w) + \lambda(1-\alpha)w + \lambda\alpha \,\mathrm{sign}(w)$$

$$\frac{\partial J}{\partial w_j} \in \frac{\partial \ell}{\partial w_j} + \lambda(1-\alpha)w_j + \lambda\alpha s_j, \quad s_j \in \mathrm{sign}(w_j)$$

Digitalizado com CamScanner

Exercício 5 (P8 3)

Com Logística Binária $\left(p = \sigma(Xw + b)\right)$.

$$\nabla \ell(w) = \frac{1}{n} X^T (p - y), \quad \frac{\partial \ell}{\partial b} = \frac{1}{n} \sum (p_i - y_i)$$

Hessiana

$$\nabla^2 J_{fusão}(w) = \nabla^2 \ell(w) + \lambda (1 - \alpha) I$$

Subgradiente direto

$$w \leftarrow w - \eta\left(\nabla \ell(w) + \lambda (1 - \alpha) w + \lambda \alpha s\right), \quad s \in \operatorname{sign}(w)$$

exercício 5 Pg 4

Logits: $z_i = x_i^T w + b$

Probabilidade: $p_i = \sigma(z_i) = \frac{1}{1+e^{-z_i}}$

Perda média (cross entropy)

$$\ell(w,b) = \frac{1}{N}\sum_{i=1}^{N}\left[-y_i \log_2 p_i - (1-y_i)\log_2(1-p_i)\right].$$

Elastic net:

$$R(w) = \lambda\left[\frac{1-\alpha}{2}\|w\|_2^2 + \alpha\|w\|_1\right], \quad \lambda \geq 0, \quad \alpha \in [0,1]$$

objetivo

$$J(w,b) = \ell(w,b) + R(w)$$

Exercício 5 (P8 5)

Derivada da perda Logística

$$l_i = -y_i \log p_i - (1-y_i)\log(1-p_i).$$

$$\frac{\partial l_i}{\partial p_i} = -\frac{y_i}{p_i} + \frac{1-y_i}{1-p_i}, \quad \frac{dp_i}{dz_i} = p_i(1-p_i).$$

$$\frac{\partial l_i}{\partial z_i} = \left(\frac{-y_i}{p_i} + \frac{1-y_i}{1-p_i}\right) p_i(1-p_i) = p_i - y_i$$

Tendo $z_i = x_i w + b$

$$\frac{\partial l}{\partial w} = \frac{1}{N}\sum_i (p_i - y_i) x_i = \frac{1}{N} X^T(p-y),$$

$$\frac{\partial l}{\partial b} = \frac{1}{N}\sum_i (p_i -$$

$$\nabla_w l = \frac{1}{N} X^T(p-y), \quad \frac{\partial l}{\partial b} = \frac{1}{N} 1^T(p-y)$$